

# Cuestionario de teoría 1

Antonio Álvarez Caballero  
[analca3@correo.ugr.es](mailto:analca3@correo.ugr.es)

28 de marzo de 2016

## 1. Cuestiones

**Cuestión 1.** Identificar, para cada una de las siguientes tareas, qué tipo de aprendizaje automático es el adecuado (supervisado, no supervisado, por refuerzo) y los datos de aprendizaje que deberíamos usar. Si una tarea se ajusta a más de un tipo, explicar cómo y describir los datos para cada tipo.

- a) Categorizar un grupo de animales vertebrados en pájaros, mamíferos, reptiles, aves y anfibios.
- b) Clasificación automática de cartas por distrito postal.
- c) Decidir si un determinado índice del mercado de valores subirá o bajará dentro de un periodo de tiempo fijado.

**Solución.** a) Aprendizaje supervisado. Lo importante aquí es el conjunto de datos, ya que cada uno de estos grupos

- b) Clasificación automática de cartas por distrito postal.
- c) Decidir si un determinado índice del mercado de valores subirá o bajará dentro de un periodo de tiempo fijado.

**Cuestión 2.** ¿Cuáles de los siguientes problemas son más adecuados para una aproximación por aprendizaje y cuáles más adecuados para una aproximación por diseño? Justificar la decisión.

- a) Determinar el ciclo óptimo para las luces de los semáforos en un cruce con mucho tráfico.
- b) Determinar los ingresos medios de una persona a partir de sus datos de nivel de educación, edad, experiencia y estatus social.
- c) Determinar si se debe aplicar una campaña de vacunación contra una enfermedad.

**Solución.**

**Cuestión 3.** Construir un problema de *aprendizaje desde datos* para un problema de selección de fruta en una explotación agraria (ver transparencias de clase). Identificar y describir cada uno de sus elementos formales. Justificar las decisiones.

**Solución.**

**Cuestión 4.** Suponga un modelo PLA y un dato  $x(t)$  mal clasificado respecto de dicho modelo. Probar que la regla de adaptación de pesos del PLA es un movimiento en la dirección correcta para clasificar bien  $x(t)$ .

**Solución.**

**Cuestión 5.** Considere el enunciado del ejercicio 2 de la sección FACTIBILIDAD DEL APRENDIZAJE de la relación de apoyo.

- a) Si  $p = 0,9$ , ¿cuál es la probabilidad de que  $S$  produzca una hipótesis mejor que  $C$ ?
- b) ¿Existe un valor de  $p$  para el cual es más probable que  $C$  produzca una hipótesis mejor que  $S$ ?

**Solución.**

**Cuestión 6.** La desigualdad de Hoeffding modificada nos da una forma de caracterizar el error de generalización con una cota probabilística

$$\mathbb{P}[|E_{out}(g) - E_{in}(g)| > \epsilon] \leq 2Me^{-2N^2\epsilon}$$

para cualquier  $\epsilon > 0$ . Si fijamos  $\epsilon = 0,05$  y queremos que la cota probabilística  $2Me^{-2N^2\epsilon}$  sea como máximo  $0,03$ , ¿cuál será el valor más pequeño de  $N$  que verifique estas condiciones si  $M = 1$ ? Repetir para  $M = 10$  y para  $M = 100$ .

**Solución.**

**Cuestión 7.** Consideremos el modelo de aprendizaje  $M$ -intervalos donde  $h : \mathbb{R} \rightarrow -1, +1$  y  $h(x) = +1$  si el punto está dentro de cualquiera de  $m$  intervalos arbitrariamente elegidos y  $-1$  en otro caso. ¿Cuál es el más pequeño punto de ruptura para este conjunto de hipótesis?

**Solución.**

**Cuestión 8.** Suponga un conjunto de  $k^*$  puntos  $x_1, x_2, \dots, x_{k^*}$  sobre los cuales la clase  $H$  implementa  $< 2^{k^*}$  dicotomías. ¿Cuáles de las siguientes afirmaciones son correctas?

- a)  $k^*$  es un punto de ruptura
- b)  $k^*$  no es un punto de ruptura
- c) todos los puntos de ruptura son estrictamente mayores que  $k^*$
- d) todos los puntos de ruptura son menores o iguales a  $k^*$
- e) no conocemos nada acerca del punto de ruptura

**Solución.**

**Cuestión 9.** Para todo conjunto de  $k^*$  puntos,  $H$  implementa  $< 2^{k^*}$  dicotomías. ¿Cuáles de las siguientes afirmaciones son correctas?

- a)  $k^*$  es un punto de ruptura
- b)  $k^*$  no es un punto de ruptura
- c) todos los  $k \geq k^*$  son puntos de ruptura
- d) todos los  $k < k^*$  son puntos de ruptura
- e) no conocemos nada acerca del punto de ruptura

**Solución.**

**Cuestión 10.** Si queremos mostrar que  $k^*$  es un punto de ruptura, ¿cuáles de las siguientes afirmaciones nos servirían para ello?:

- a) Mostrar que existe un conjunto de  $k^*$  puntos  $x_1, \dots, x_{k^*}$  que  $H$  puede separar (*shatter*).
- b) Mostrar que  $H$  puede separar cualquier conjunto de  $k^*$  puntos.
- c) Mostrar un conjunto de  $k^*$  puntos  $x_1, \dots, x_{k^*}$  que  $H$  no puede separar.
- d) Mostrar que  $H$  no puede separar ningún conjunto de  $k^*$  puntos.
- e) Mostrar que  $m_H(k) = 2^{k^*}$

**Solución.**

**Cuestión 11.** Para un conjunto  $H$  con  $d_{VC} = 10$ , ¿qué tamaño muestral se necesita (según la cota de generalización) para tener un 95 % de confianza de que el error de generalización sea como mucho 0,05?

**Solución.**

**Cuestión 12.** Consideremos un escenario de aprendizaje simple. Supongamos que la dimensión de entrada es uno. Supongamos que la variable de entrada  $x$  está uniformemente distribuida en el intervalo  $[-1, 1]$  y el conjunto de datos consiste en 2 puntos  $x_1, x_2$  y que la función objetivo es  $f(x) = x^2$ . Por tanto el conjunto de datos completo es  $\mathcal{D} = (x_1, x_1^2), (x_2, x_2^2)$ . El algoritmo de aprendizaje devuelve la línea que ajusta estos dos puntos como  $g$  (i.e.  $H$  consiste en funciones de la forma  $h(x) = ax + b$ ).

- a) Dar una expresión analítica para la función promedio  $\bar{g}(x)$ .
- b) Calcular analíticamente los valores de  $E_{out}$ , *bias* y *var*.

**Solución.**

**BONUS.** Considere el enunciado del ejercicio 2 de la sección ERROR Y RUIDO de la relación de apoyo.

- a) Si su algoritmo busca la hipótesis  $h$  que minimiza la suma de los valores absolutos de los errores de la muestra,

$$E_{in}(h) = \sum_{n=1}^N |h - y_n|$$

entonces mostrar que la estimación será la mediana de la muestra,  $h_{med}$  (cualquier valor que deje la mitad de la muestra a su derecha y la mitad a su izquierda)

- b) Suponga que  $y_n$  es modificado como  $y_n + \epsilon$ , donde  $\epsilon \rightarrow \infty$ . Obviamente el valor de  $y_n$  se convierte en un punto muy alejado de su valor original. ¿Cómo afecta esto a los estimadores dados por  $h_{mean}$  y  $h_{med}$ ?

**Solución.**

**BONUS.** Considere el ejercicio 12.

- a) Describir un experimento que podamos ejecutar para determinar (numéricamente)  $g(x)$ ,  $E_{out}$ , *bias* y *var*.
- b) Ejecutar el experimento y dar los resultados. Comparar  $E_{out}$  con *bias+var*. Dibujar en unos mismos ejes  $g(x)$ ,  $E_{out}$  y  $f(x)$ .

**Solución.**